悪条件問題に対する数値解法(数値計算のアルゴリズムの研究)

AUTHOR(S):

北川, 高嗣

CITATION:

北川, 高嗣. 悪条件問題に対する数値解法(数値計算のアルゴリズムの研究). 数理解析研究所講究録 1983, 483: 223-237

ISSUE DATE:

1983-03

URL:

http://hdl.handle.net/2433/103419

数理解析研究所講究録
483巻 1983年 223-237

悪条件問題に対する数値解法

愛媛大学 理学部 北川高嗣
(Takashi Kitagawa)

本稿に於いては、悪条件問題 (Ill-posed Problems) に対する数値的取り扱いについて考える。ここで扱う悪条件問題とは、以下の様なものを指すこととする。

1. 線型作用素方程式に於けるIll-posednessの特徴づけ。

$$Kf = g \quad : \quad K \in [X \to Y],\ f \in X,\ g \in Y.$$

ここにKは、完備距離空間XからYへのコンパクト線型作用素とし、f, g, は各々X, Yの要素とする。

Hadamardによって定義されたWell-posedの概念は、次の二つの条件を満たすものとして知られている。

a). Yの任意の要素 $g \in Y$ に対し解fがX内に一意に存在する。(i.e. for $\forall g \in Y$, $\exists f \in X$ s.t. $Kf = g$.)

b) 解 $f \in X$ は、gに連続に依存する。

(The solution $f \in X$ depends continuously on $g \in Y$)

Well-posedでない問題を、Ill-posed Problemsという。

b) の条件は安定性(Stability)に関するものであるが、厳密に述べれば次の様になる。

今 $Kf=g$, $K\in[X\to Y]$: compact, linear $X;Y$: complete Metric sp. であるが、X,Y の metric を各々 $\rho_X(\cdot,\cdot)$, $\rho_Y(\cdot,\cdot)$ とすれば、

For $^{\forall}\varepsilon>0$ $^{\exists}\delta(\varepsilon)>0$ such that

$$\rho_Y(g_1,g_2)\leq\delta(\varepsilon)\Rightarrow\rho_X(f_1,f_2)\leq\varepsilon$$

where $Kf_1=g_1$, $Kf_2=g_2$, $f_1,f_2\in X$, $g_1,g_2\in Y$.

この条件は線型作用素Kに関して言えば、

「Kの逆作用素(Inverse operator) K^{-1} が不連続。」

である事を意味し、これは「K^{-1} が非有界。」と同値である。

悪条件作用素方程式の性質は、特異関数(Singular Function)で展開する事により、容易に理解される。今Kの特異関数を ϕ_r, ψ_r, それに対応する特異値を λ_r とすれば、

$$K\psi_r=\lambda_r\phi_r,\ \text{かつ}\ K^*\phi_r=\lambda_r\psi_r,\quad\text{或いは、}$$

$$KK^*\phi_r=\lambda_r^2\phi_r,\ K^*K\psi_r=\lambda_r^2\psi_r\quad(r=0,1,\cdots\cdots)$$

である。今 $X=Y=L_2(a,b)$ ととれば、ϕ_r,ψ_r は orthonormal system であり $L_2(a,b)$ の basis となっている。このとき、

$\lambda_0^2>\lambda_1^2>\cdots\cdots>\lambda_n^2>\cdots\cdots\to0$ と singular values の列は $n\to\infty$ で $\lambda_n^2\to0$ に収束する。(Ill-posed problems の上記の性質「K^{-1} 非有界」は $\lambda_n^2\to0$, as $n\to\infty$ によって特徴づけられる。)

ここで元の方程式 $Kf=g$ の解 f を Singular functions と Singular Values を使って表わせば以下の様になる。

$$f = \sum_{i=1}^{\infty} \frac{1}{\lambda_i} \langle g, \phi_i \rangle \psi_i \qquad (1)$$

$\phi_i,\ \psi_i$: Singular functions $\quad i=1,2,\cdots\cdots$

λ_i : Singular values $\quad i=1,2,\cdots\cdots$

where $\lambda_1^2 > \lambda_2^2 > \cdots > \lambda_n^2 > \cdots \to 0$

2. 解 f の存在と一意性.

前節の(1)式から、ただちに次の結果が得られる。

定理1 (Picard)

前節の線型作用素方程式が解をもつ為の必要十分条件は. (i) $\sum_{i=1}^{\infty} \frac{1}{\lambda_i^2} |\langle g, \phi_i \rangle|^2 < \infty$

かつ (ii) $g \in R(K)$ なる事である。

一意性に関しては. 次の事が言える。今 $\{\lambda_i\}$ は 0 に収束してゆく sequence である。今 もし $\lambda_i = 0$ $i \geq^{\exists} k$ であれば解に一意性はない。よって解が一意である為には. λ_i は limit point での 0 となる列でなければならない。(即ち $\lambda_i \neq 0$)

具体的に悪条件線型作用素方程式の数値的取り扱いを考える時、一つの重要な要素は $\lambda_i \to 0$ へ収束する速さである。

λ_i の収束の速さにより方程式の"conditioning"が支配される。

3. 第一種積分方程式の"conditioning".

今、具体的に線型作用素 K として、次を考える。

$$Kf = \int_a^b k(t,s) f(s)\, ds \qquad (2)$$

この時 方程式 $Kf = g$ は、第一種積分方程式（フレドホルムタイプ）となる。多くの重要な境界条件問題（数学的、物理的諸問題）が、この第一種積分方程式に帰着される事が知られている。今 $f, g \in L_2(a,b)$ とし $k(s,t)$ は、L_2 カーネル $\int_a^b \int_a^b |k(s,t)|^2 dt\, ds < \infty$ とする。

この時、作用素 K の Singular values の収束に関し、次の定理がある。

定理 2 (H. Weyl)

今、$k(t,s) = k(s,t)$ とする。（この場合 Singular values は Eigen values に等しい。）

$k(t,s)$ が p 階微分可能でかつ p 次導関数が連続であれば、$|\lambda_n| = O(n^{-p-\frac{1}{2}})$.

$$\left(\text{i.e. } \begin{array}{l} k(t,s) = k(s,t) \\ \exists\ \dfrac{\partial^r k(t,s)}{\partial t^r} \in C^0 \quad \text{for } r = 0, \cdots, p \end{array} \Rightarrow |\lambda_n| = O(n^{-p-\frac{1}{2}}). \right)$$

さらに、$k(s,t) \neq k(t,s)$ の場合に対し、より精密化された

結果として次がある。

定理3 （F. Smithies ）

$1<\beta<2$ とする。ある $p\geq 0$ に対し $k(t,s)$ が，

(i) $k^{(r)}(t,s)\equiv\dfrac{\partial^r k(t,s)}{\partial t^r}$ が存在し，（ほとんどすべての s と $r=0,\cdots,n-1$ に対し），$k^{(n)}(t,s)$ が 絶対連続（in t），

(ii) $k^{(p)}(t,s)\in L_\beta$ （ $p>0,\ 1<\beta<2$ ）

(iii) $\displaystyle\int_a^b\left\{\int_a^b|k^{(p)}(t+\theta,s)-k^{(p)}(t-\theta,s)|^\beta dt\right\}^{2/\beta}ds\leq A|\theta|^{2\alpha}$ for some A and all suff. small θ,

(iv) $p>0$，$\alpha>0$ かつは $p=0$ ，$\alpha>\dfrac{1}{\beta}-\dfrac{1}{2}$

である時.

$$\lambda_n=O(n^{-\alpha-1-p+1/\beta})。$$

定理2，定理3．の結論は、共に．積分作用素のカーネル $k(s,t)$ が 滑らかであればある程，Singular value の 0 に収束する速さが増す事を意味する。よく知られた例として．$k(x,y)$ が，対数特異点を持つ様な場合，積分作用素 K の Singular values は $O(\frac{1}{n})$ 程度でしか収束しない。この様な場合は．積分の実行は、困難であるが，"conditioning" の点から言えば．右辺関数 f の精度さえ十分にあれば．ほとんど問題にはならない。これに対し，カーネル $k(x,y)$ が十分滑らか．(Analytic) な場合一般に，作用素 K の特異値は 急速に（exponential に） 0 に収

束してゆくと考えられる。この様な場合とは、どんな方法で離散化しても、直接、離散化された線型方程式を解く事によって満足る解を得る事は出来ない。(nを十分大きくとった場合。) 直接法による数値解のもとめにくさはさておき(この問題点を避ける数値解法については後に体系的に述べる。)意味のある解がもとまるための必要条件は、$(g.\varphi_i)$ が十分速く収束する事である。ここで言う十分速くの意味は、λ_iが0に収束するよりも速く、という意味である。この条件は、(2)のカーネルが前述した様に、解析的である場合には、λ_iは0に相当速く、指数的に収束する為、gに対する強い制約となる。しかしながらgがその制約を満たさなければ、意味のある解、増してや数値解は、もとめる事は出来ない。

'悪条件問題の解きにくさ'は、確かに広く知られている様に、系の'conditioning'に依存する。それは、直接法のみを考える時、$\lambda_i \to 0$ の収束の速さのみに依存する様に見える。 しかし、後述する方法群を用いれば、$\lambda_i \to 0$、つまり'conditioning'自体は、重要な問題とはならない。

例えば、次の様な場合を考えてみれば、この事は、容易に理解されるであろう。 今、カーネルが、解析的であり、λ_iが指数的に0に収束すると仮定する。(実際、問題となるのは、この様な場合である。) つまり、離散化された時の、

線型連立方程式の，いわゆる条件数（Condition Number）が，指数的に増加してゆく様な場合である。この場合，直接的な解法による解は，一般に全く信頼できない。ところが，今，(g, ϕ_i) が，非常に速く収束する，或いは，小さな i にのみ対し非零であると仮定してみる。例えば，$(g, \phi_i)=0$ $i \geq k$ としてみよう。この時，λ_i が $i>k$ において，いくら速く収束しようとも，単に，解 f，即ち

$$f=\sum_{i=1}^{\infty} \frac{1}{\lambda_i}(g, \phi_i)\psi_i$$

において，k 項以後を，強制的に 0 と置く事により，十分満足できる解，（従って最値解）をもとめる事が出来る。

これが，悪条件問題に対する数値解法を考える時の一つの拠り所である。以下に，4つの代表的な悪条件問題に対する数値解法を列挙し，上記の視点から，解法間の関係について見直してみる。

3．悪条件問題に対する数値解法とその解法間の関係

3.a) 最小自乗最小ノルム解．（Minimum Norm Least Squares Solution.）

M.N.L.S.S と略される。この解を f_a とすれば，

$$f_a=\{f \mid \min_{f} \|Kf-g\|^2\}$$

とし，（即ち，最小自乗解に一意性がなくその集合を f_a とし

た時。この時 $\mathcal{N}(K) \neq \{0\}$, $\mathcal{N}(K)$: Null Space of K)

$$f_a \in \min_{f \in B_a} \|f\|^2$$

この場合，Singular values に関して言えば，$\exists k$ s.t. $\lambda_i = 0$ for $i > k$ である事を意味する。$\lambda_i = 0$ に対する，フーリエ係数 (g, ϕ_i) と $\frac{1}{\lambda_i}$ の積は，強制的に，0 と置く。

3,b) Truncated Singular (eigen) Value (T.S.V.)

T.S.V. と以下略す。その解を f_b，とすれば，

$$B_b = \mathrm{Span}(\psi_1, \psi_2, \cdots, \psi_n)$$

$$f_b \in \min_{f \in B_b} \|Kf - g\|^2$$

特異関数で表現した場合には，以下の様に表現される。

$$f_b = \sum_{i=1}^{n} \frac{1}{\lambda_i} (g, \phi_i) \psi_i$$

一般に $\lambda_i \neq 0$ $i > n$ である。微小な λ_i による誤差の増大が誘発される前に，n で Series を打ち切ってしまう。

3.c) Standard Regularization (標準型正則化法)

S.R と略す事にする。この解を f_c，とすれば，

$$f_c \in \min_f \left(\|Kf - g\|^2 + \mu \|f\|^2 \right)$$

ここに，μ は Regularization parameter と呼ばれる。
又これは，方程式

$$K^*Kf + \mu f = K^*g$$

を解いて得られる解と同等である。f_c，を陽に表わせば，

$$f_c = (K^*K + I\mu)^{-1} K^*g$$

又，特異関数系によって表わせば

$$f_c = \sum_{i=1}^{\infty} \frac{\lambda_i}{\lambda_i^2 + \mu} (g, \phi_i) \psi_i$$

となる。

3.d).（Modified）Regularization.

一般に正則化法と言えばこれを指す。M.R. と略す事にし，その解を f_d とすれば，

$$f_d \in \min_f \left(\|Kf - g\|^2 + \mu \|\Omega f\|^2 \right)$$

ここに μ は Regularization Parameter，Ω は Stabilizer（一般に Ω は differential operator）と呼ばれる。f_d を陽に表わせば，

$$f_d = (K^*K + \mu\Omega^*\Omega)^{-1} K^*g$$

と表わされる。以下に方法間の関係について述べる。

i) 3.a) M.N.L.S.S (Minimum Norm Least Squares Solution) と
3.b) T.S.V (Truncated Singular Value) の関係.

3.a) と 3.b) は形式上は全く同じに見える。しかし解の意味は厳密に区別されるべきである。

M.N.L.S.S. は $\lambda_i \neq 0$ に対応する特異関数全てを使って解を構成する。T.S.V. は $\lambda_i \neq 0$ に対する特異関数の一部（選び方には任意性が残る。）を使って解を構成する。

つまり M.N.L.S.S に関しては構成する特異関数の選び方にあいまいさはない。T.S.V に於いては解は特異関数の部分空間の選び方に依存する。

ii) 3.c) Standard Regularization (S.R) と
3.d) Modified Regularization (M.R) の関係。

3.d) の M.R. は 3.c) の S.R の一般化になっている。即ち M.R. に於いて Stabilizer $L = I$ とおけば S.R と同値となる。一般に、M.R $L \neq I$ は S.R. $L = I$ の場合に帰着する事が出来る。具体的には L が invertible ($\exists L^{-1}$) であれば Cholesky 分解によって、L が invertible でなければ ($\nexists L^{-1}$) Q−R 分解によって標準型正則化法 ($L = I$) に帰着する事ができる。S.R、M.R とも解は Regularization Parameter μ の選び方に依存する。3.c) は又 Damped Least Squares と呼ばれる事もある。

iii) 3b) T.S.V. と 3c) Regularization の間の関係。

3b) の解 f_b，3c の解 f_c を，特異関数系を使って書いてみると，

3.b) T.S.V. $\{f_b\}_i = \begin{cases} \dfrac{1}{\lambda_i}(g, \phi_i)\psi_i & i \le n \\ 0 & i > n \end{cases}$

3.c) S.R $\{f_c\}_i = \dfrac{\lambda_i}{\lambda_i^2 + \mu}(g, \phi_i)\psi_i \quad i = 1, 2, \cdots\cdots$

となり，どちらも同じ特異関数で解を構成し，同じフーリエ係数成分 (g, ϕ_i) を含んでいる。3.b) と 3.c) の形式上の違いは，わずかに，最初の Singular value を含む係数が

T.S.V. では $\begin{cases} \dfrac{1}{\lambda_i} & (i \le n) \\ 0 & (i > n) \end{cases}$　S.R では $\dfrac{\lambda_i}{\lambda_i^2 + \mu}$　$i = 1, 2, \cdots$

となっている所である。これらを Damping factor と呼ぶ事にしよう。

ここで，作用素 K の Singular value が，指数的に 0 に収束してゆく場合，(例えば，第一種積分方程式に於いて，カーネル $k(s,t)$ が Analytic な場合) には，Regularization parameter μ に，適当な変数変換を行なってやれば，T.S.V と S.R は，ほぼ同一視する事が出来ることを示す。

まず，Standard Regularization の Damping factor を $d(\mu)$ とおく。即ち $d^i(\mu) = \lambda_i / (\lambda_i^2 + \mu)$.

今、Kの singular values λ_i が、指数的に0に収束してゆくと仮定する。その速さが β^{-i} に比例するとすれば、$(\beta \gg 1)$

$\lambda_i = O(\beta^{-i})$ である。ここで、Regularization parameter μ、に対し $\nu = -\log_\beta \mu$ なる変数変換を行なう。これは regularization parameter μ を、λ_i の収束に合わせ指数的に変化させる事を意味する。つまり $\mu = \beta^{-\nu}$ である。

ここで $\mu = \beta^{-\nu}$ が $\lambda_m^2 > \mu > \lambda_{m+1}^2$ にある場合を考えれば、

$$d^i(\mu) = \frac{\lambda_i}{\lambda_i^2 + \mu} = \frac{1/\lambda_i}{1 + \frac{\mu}{\lambda_i^2}} = \frac{1/\lambda_i}{1 + \frac{\beta^{-\nu}}{\lambda_i^2}}$$

今 $i \leq m$ とすれば、$\beta^{-\nu} = o(\lambda_i^2)$

つまり、 $d^i(\mu) \simeq \dfrac{1}{\lambda_i}$ for $i \leq m$

($\simeq$ の近さは、β の大きさに依存する。)

さらに $i > m$ とすれば、逆に $\beta^{-\nu} \gg \lambda_i^2$、$i > m$ であるから $d^i(\mu) \simeq 0$ for $i > m$。

まとめて書けば、

$$d^i(\mu) \simeq \begin{cases} \dfrac{1}{\lambda_i} & i \leq m \\ 0 & i > m \end{cases}$$

となり、これは Truncated Singular Value の Damping factor に一致する。この傾向は、β が大きければ大きい程、即ち、λ_i が0に収束する速さが、速ければ速い程、強くなる。

さらに，Damping factor の比 $R^i(\mu)$ を

$$R^i(\mu) = \frac{\lambda_i}{\lambda_i^2+\mu} \Big/ \frac{1}{\lambda_i} \qquad (\mu = \beta^{-\nu})$$

$$= \left(1 + \frac{\beta^{-\nu}}{\lambda_i^2}\right)^{-1}$$

とおけば，

$$R^i(\mu) \simeq 1 \qquad \lambda_i^2 \gg \beta^{-\nu}$$

$$= 0 \qquad \lambda_i^2 \ll \beta^{-\nu}$$

となり，λ_i の収束が速い事（即ち $\beta \gg 1$），$\nu = -\log_\beta \mu$ の変数変換を行なう事，の二つの条件の下で，Truncated Singular Value のしきい値 τ，（しきい値の意味は，$\lambda_i \geqq \tau$ であれば，$\frac{1}{\lambda_i}$，$\lambda_i < \tau$ であれば，0 とする判定値）と Regularization parameter μ の間には，$\tau \simeq \sqrt{\mu}$ なる関係が~~成~~立する事がわかる。

以上により，既存の数値解法に対し，一つの統一的な視点が与えられた。これは，例えば，パラメータ μ を決定する時等に有用である。一つの数値解法で，（今，S.R で）μ の値が決まったとし，その値を μ_{opt} とすれば，それは同時に，T.S.V. の最適しきい値 τ_{opt} を $\tau_{opt} = \sqrt{\mu_{opt}}$ で与える事になる。

References.

[1] Smithies, F. "The eigenvalues and singular values of integral equations", Proc. London Math. Soc., 43 (1937), pp 255-279.

[2] Weyl, H. Das asymptotische verteilungsgestz der eigenwerte lineare partieller differentialgleichungen, Math. Ananlen., 71 (1912), pp 441-479.

[3] Lewis, B.A. On the numerical solution of Fredholm integral equations of the 1st kind, JIMA, 16 (1975), pp 207-220.

[4] Varah, J.M. On the numerical solution of ill-conditioned linear systems with application to ill-posed problems, SIAM J. Num. Anal. Vol. 10, No. 2, April 1973. pp 257 - 266.

[5] ———— A practical examination of some numerical methods for linear descrete ill-posed problems, SIAM Rev. Vol 21 No. 1 (1979)

[6] Wahba, G. Practical approximate solutions to linear operator equations when the data are noisy. SIAM J. Num. Anal Vol 14 No. 4 (1977) pp 651-667.

[7] Hilgers, J.W. On the equivalence of regularization and ~~certain~~ reproducing Kernel Hilbert space approaches for solving first kind problems.

SIAM J. Num. Anal., 13, (1976) pp 172-184.

[8] Hanson, R. A numerical method for solving Fredholm integral equations of the first kind using singular values. SIAM J Num. Anal. 8. 1971, pp 616-622.

[9] Tikhonov, A.N. Solutions of ill-posed problems. 1977. V.H. Winston & Sons.